

EX PARTE OR LATE FILED

ORIGINAL

U S WEST, Inc.  
1801 California Street, Suite 5100  
Denver, Colorado 80202  
303 672-2842  
Facsimile 303 296-4576

**USWEST**

**Vincent C. DeGarlais**  
Senior Attorney

RECEIVED

JUL 1 1999

FEDERAL COMMUNICATIONS COMMISSION  
OFFICE OF THE SECRETARY

July 1, 1999

Ms. Magalie Roman Salas, Secretary  
Federal Communications Commission  
The Portals  
445 Twelfth Street S.W.  
Washington, D.C. 20554

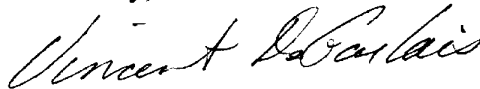
RE: CC Docket No. 98-121

Dear Ms. Salas:

On Tuesday June 29, Jeff Owens, Michael Williams, Michael Carnall, John Devaney and myself, representing U S WEST, met with David Kirschner, Eric Einhorn, Florence Setzer, Daniel Shiman, Alex Belinfante, John Stanley, Michael Pryor, and Claudia Pabo of the Common Carrier Bureau staff to discuss U S WEST's performance measurement plan. Copies of the materials discussed in the meeting are attached.

In accordance with Section 1.1206(b)(2) of the Commission's rules, an original and one copy of this letter and the attachments are being filed with your office for inclusion in the record of this proceeding. All aforementioned Commission staff will also be served a copy of this ex parte.

Sincerely,



Vincent C. DeGarlais

VCD/jvv  
Enc.

No. of Copies rec'd 41  
List ABCDE

# **SELF-EXECUTING PENALTY PLANS: SOME FUNDAMENTAL CONCERNS**

- Self-Executing plans raise due process concerns
- Self-Executing plans conflict with Department of Justice policy and judicial decisions
- Self-Executing plans rely upon unsound statistical inferences
- Self-Executing plans encourage CLEC inefficiency and fail to identify the causes for statistical disparities

# THE REQUIREMENTS OF DUE PROCESS

- Statistical evidence is always rebuttable evidence
- The respondent must have a chance to respond -- i.e., to explain or rebut the evidence
- No penalty should be imposed unless the rebuttal evidence is first considered

# **THE DEPARTMENT OF JUSTICE AND THE COURTS HAVE ESTABLISHED THAT STATISTICS ARE REBUTTABLE**

- Employment discrimination provides a relevant comparison
- In briefs the Department of Justice has emphasized the rebuttable nature of statistical evidence
- Courts, including the Supreme Court, have established that statistics are rebuttable
- Self-executing penalty plans are inconsistent with DOJ position and judicial precedent

# **REASONS FOR NOT TREATING STATISTICAL DISPARITIES AS DISPOSITIVE PROOF OF DISCRIMINATION**

- Statistical disparities often are attributable to nondiscriminatory causes
- Statistical disparities can result from data errors
- Treating statistical disparities as dispositive proof of discrimination may encourage inefficient conduct by CLECs intended to cause disparities

# CONCLUSIONS

- Statistics are important in analyzing potential discrimination, but the reasons for statistical disparities must be considered to determine whether discrimination has occurred
- The Act's purpose of encouraging competition supports determining the reasons for statistical disparities
- To ensure compliance with due process and to encourage fair competition, it is necessary to consider the cause of statistical disparities.

**THE USE OF STATISTICS IN  
SELF-EXECUTING PENALTY PLANS  
AND APPLICATIONS UNDER SECTION 271  
OF THE TELECOMMUNICATIONS ACT OF 1996**

**I. Introduction**

The FCC suggests that "self-executing" penalty plans may be an essential part of any acceptable proposal for compliance with the requirements of section 271 of the Telecommunications Act of 1996 (the "Act"). However, a self-executing penalty plan that would impose penalties upon a Bell operating company ("BOC") simply by virtue of bare statistical anomalies or disparities in performance measures would violate a BOC's constitutional due process rights. Due process requires that, prior to the imposition of any monetary penalties, the entity penalized must have an opportunity to respond to allegations of discrimination. This means that self-executing penalty plans by their very nature violate due process.

Where allegations of discrimination are based upon conclusions drawn from statistical data, due process requires that the charged party be given the opportunity to rebut the purported statistical proof or to explain the apparent statistical disparity. This principle should apply with the same force in section 271 proceedings as it does in the employment discrimination cases that gave rise to the principle. In short, self-executing penalty plans are unlawful because they fail to guarantee the right to respond to charges of discrimination based upon bare statistical disparities.

In addition to being inconsistent with the requirements of due process, self-executing penalty plans are not grounded in sound public policy. As demonstrated by more than 30 years of reported judicial decisions in the area of employment law, statistical disparities often are attributable to reasons unrelated to discriminatory treatment. It is in the public interest to determine the true reasons for statistical disparities in the provisioning of telecommunications services. That interest is not served by a policy that unrealistically assumes any statistical disparity is caused by discrimination and cannot be explained by other factors.

In the context of employment discrimination cases, the United States Department of Justice ("DOJ") has consistently recognized the importance of giving parties the opportunity to explain statistical disparities. The DOJ has urged the courts that respondents must be given the opportunity to respond to allegations of

discrimination stemming from bare statistical anomalies. As shown below, the courts have uniformly adopted this principle.

There is no legitimate reason why this principle should not apply with equal force when analyzing possible discrimination in the context of a self-executing penalty plan or an application under section 271. Disparities in statistical data, in and of themselves, do not show that a BOC has discriminated against its competitors. The disparities could be explained by a variety of factors unrelated to discriminatory service, including, for example, weather, location of customers, an usually high volume of orders, and the time of the month or year. All of these nondiscriminatory factors can affect the time a BOC requires to perform certain tasks and operations. For example, if a CLEC places a large percentage of its orders toward the end of the month, the time required to complete the orders could be greater than the time required for orders submitted at the beginning of the month.

These are the types of factors that are overlooked if the BOC is not given the opportunity to respond to allegations of discrimination based on bare statistical disparities. It is essential that any analysis of potential discrimination based on statistical results allow for consideration of these factors. If potential nondiscriminatory explanations are excluded from an analytical framework, false findings of discrimination may result.

This paper addresses the proper role of inferences from statistical data in section 271 proceedings. Because statistical data are always rebuttable, both due process and sound public policy require that BOCs have the opportunity to respond to statistical disparities before the FCC (or any FCC-sanctioned penalty plan) imposes monetary penalties. For this reason, any acceptable penalty plan must give BOCs the opportunity to rebut or explain any adverse inferences drawn from bare statistical disparities.

## **II. The Role of Statistical Evidence Within the FCC's Analytical Framework**

The FCC has demonstrated it will rely heavily on statistical evidence in evaluating the issues that arise under section 271. First, the FCC uses the statistical performance data to analyze a BOC's compliance with the fourteen-point competitive checklist. Second, the FCC intends to use statistical data gathered after granting section 271 applications to ensure that BOCs continue to keep their local telecommunications markets open to competition.

The FCC's analytical framework for evaluating a section 271 application requires a BOC to make a prima facie showing that it meets the requirements of each

of the 14 checklist items. The FCC has repeatedly emphasized the importance of performance data in analyzing whether a BOC has made a showing that it is providing nondiscriminatory service, giving CLECs a meaningful opportunity to compete, and satisfying the competitive checklist in section 271. See, e.g., Application of BellSouth Corporation et al. for Provision of In-Region InterLATA Services in Louisiana, Memorandum Opinion and Order, CC Docket No. 98-121, FCC 98-271, at ¶ 72 (rel. Oct. 13, 1998) ("BellSouth Louisiana II Order"); Application of Ameritech Michigan Pursuant to Section 271 of the Communications Act of 1934, as amended, to Provide In-Region, InterLATA Service in Michigan, CC Docket No. 97-137, Memorandum Opinion and Order ¶ 141 (rel. Aug. 19, 1997) ("Ameritech Michigan Order").

The FCC's decisions relating to section 271 applications are replete with analyses of BOCs' statistical data relating to the critical areas of performance, including interconnection, access to network elements on an unbundled basis, interim number portability ("INP"), and resale services. See, e.g., Ameritech Michigan Order ¶¶ 141, 183, 240; Application of BellSouth Corporation Pursuant to Section 271 of the Communications Act of 1934, as amended, to Provide In-Region, InterLATA Service in South Carolina, CC Docket No. 97-208, Memorandum Opinion and Order ¶ 107 (rel. Dec. 24, 1997) ("BellSouth South Carolina Order"); BellSouth Louisiana II Order ¶¶ 278-282.

The FCC's discussion of performance data in its section 271 decisions demonstrates that the agency gives substantial weight to BOC performance data showing statistical disparities in treatment, and that the weight it gives this evidence, in some cases, amounts to an irrebuttable presumption of discrimination and noncompliance with checklist items. For example, in addressing flow-through rates for operational support systems ("OSSs") in the BellSouth Louisiana II Order, the FCC gave dispositive effect to statistical disparities: "we determined that the 'substantial disparity between the flow-through rates of BellSouth's orders and those of competing carriers, on its face, demonstrates a lack of parity.'" (quoting BellSouth South Carolina Order ¶ 107); see also Application of BellSouth Corporation et al. Pursuant to Section 271 of the Communications Act of 1935, as Amended, to Provide In-Region InterLATA Services in Louisiana, Memorandum Opinion and Order, CC Docket No. 97-231, FCC 98-17 at ¶ 25 (rel. Feb. 4, 1998) (hereinafter "BellSouth Louisiana Order").

Similarly, in the BellSouth Louisiana II Order, the FCC determined that BellSouth did not provide equivalent OSS access based upon statistical disparities between the average installation intervals for CLECs and BellSouth itself. BellSouth Louisiana II Order ¶ 126. It reached a similar conclusion with respect to access to maintenance and repair OSS functions based upon statistical disparities between

dispatch times, timely repairs, and repeat trouble reports for CLECs and BellSouth. Id. ¶ 147.

Moreover, in the Ameritech Michigan Order, the FCC determined that Ameritech did not provide equivalent interconnection facilities based upon statistical disparities in trunk blocking between trunks serving Ameritech's retail operations and Ameritech's end-office interconnection trunks serving CLECs. See Ameritech Michigan Order ¶ 240; BellSouth Louisiana II Order ¶ 77 (same).

In considering post-approval compliance with section 271, the FCC has expressed its intention to give similar weight to statistical disparities. Thus, the FCC has endorsed self-executing penalty plans that would impose financial penalties on BOCs based solely on statistical disparities in performance results.

However, as the Supreme Court and the DOJ have recognized, while statistically significant disparities in the treatment of one group as compared with another can provide meaningful evidence, disparities alone should never serve as conclusive proof of discrimination. Because statistical disparities often result from factors unrelated to discrimination, both the Supreme Court and the DOJ have recognized that a party against whom discrimination has been alleged should have the opportunity to rebut or explain statistically significant differences in treatment by offering alternative explanations of the disparities.

### **III. Historical Uses of Statistical Data in Discrimination Cases**

As noted above, the dispositive effect that the FCC gives statistical disparities in performance data is inconsistent with the position historically taken by the DOJ and federal courts in employment discrimination cases. The DOJ and the courts alike have long held the view that, although statistical evidence of disparities in treatment is probative, it creates no more than rebuttable evidence of discrimination.

The DOJ's views relating to the appropriate use of statistics are particularly relevant in the context of section 271 applications, since the FCC accords substantial weight to the DOJ's evaluation of section 271 applications. Ameritech Michigan Order ¶ 37. While there obviously are substantive differences between discrimination in providing telecommunications service and discrimination in employment against a person or group of people based on race, sex, or national origin, there nevertheless is no principled reason why the DOJ's position on the use of statistics should differ depending on the type of discrimination at issue.

Moreover, a BOC's right to respond to statistical disparities should not be affected by the fact that the disparities will arise from data that the BOC itself

collected and reported. The due process concern that requires an opportunity to respond turns not on who presents the data, but, rather, on whether there is an opportunity to respond to statistical disparities.<sup>1</sup>

The DOJ has unequivocally endorsed the right to respond to showings of statistical disparities in employment discrimination cases. For example, in Hazelwood Sch. Dist. v. United States, the DOJ acknowledged that "[o]nce evidence of a substantial statistical disparity is presented, the defendant has the right to demonstrate that the disparity is not the result of discriminatory actions, but has been caused by other, racially neutral factors." Brief for the United States at 21-22; Hazelwood Sch. Dist. v. United States, 433 U.S. 299 (1977) (No. 76-255) (emphasis added).<sup>2</sup>

The Supreme Court likewise has strongly endorsed the rebuttable nature of statistical disparities in cases involving discrimination:

[O]ur cases make it unmistakably clear that "[s]tatistical analyses have served and will continue to serve an important role" in cases in which the existence of discrimination is a disputed issue . . . . We caution only that statistics are not irrefutable; they come in infinite varieties and, like any other kind of evidence, they may be rebutted. In short, their usefulness depends on all of the surrounding facts and circumstances . . . .

International Bhd. of Teamsters v. United States, 431 U.S. 324, 339-40 (1977).

Accordingly, neither courts nor parties are entitled to "assume" that a party's statistical evidence is reliable. When presented with statistical evidence of alleged discrimination, an adversary is always free to challenge the evidence by impeaching its reliability, offering rebuttal evidence, or arguing against the probative weight of the alleged statistical proof. Dothard v. Rawlinson, 433 U.S. 321, 338-39 (1977)

---

<sup>1</sup> The general need for rebuttal as a constraint on the use of statistics was emphasized by the Court in McCleskey v. Kemp, 481 U.S. 279 (1987). In McCleskey, the Court considered a death row inmate's claim that a statistical study demonstrating pervasive racial bias in the state's application of the death penalty required his conviction to be overturned. Id. The Court contrasted the inmate's claim with other cases in which it had accepted statistics as the basis of a prima facie case of discrimination, noting that " . . . Here, the State has no practical opportunity to rebut the . . . study." Id. at 296.

<sup>2</sup> The right to rebut exists in both disparate impact and disparate treatment cases: "There are prima facie and defense components to both disparate treatment and disparate impact analyses." Brief for the United States as Amicus Curiae at 17 n.17, Watson v. Fort Worth Bank & Trust, 487 U.S. 977 (1988) (No. 86-6139).

(Rehnquist, J., concurring); see also Hazelwood Sch. Dist. v. United States, 433 U.S. 299, 309-13 (1977) (appellate court committed error by disregarding evidence that could rebut proffered statistical proof).

Indeed, because of the malleability of statistics, courts examine them carefully to ensure their validity:

[S]tatistical evidence is circumstantial in character and its acceptability depends on the magnitude of the disparity it reflects, the relevance of its supporting data, and other circumstances in the case supportive of or in rebuttal of a hypothesis of discrimination. And, in reviewing statistical evidence and its supporting data, the Court must give consideration and evaluate fairly such conflicting opinions and hypotheses as may have been presented, tempering its conclusions with what one Court has described as "a pinch of common sense."

EEOC v. Federal Reserve Bank of Richmond, 698 F.2d 633, 646-47 (4th Cir. 1983), rev'd on other grounds, 467 U.S. 867 (1984). See also EEOC v. Western Elec. Co., 713 F.2d 1011, 1019 (4th Cir. 1983) (statistics must not be accepted uncritically). Consistent with these principles, courts reject statistical analyses that fail to consider neutral, nondiscriminatory explanations. Doan v. Seagate Tech., Inc., 82 F.3d 974, 979 (10th Cir. 1996), cert. denied, 519 U.S. 1056 (1997) (analysis did not account for the difference in segments of the workforce, which tended to discount plaintiff's claim); Contreras v. City of Los Angeles, 656 F.2d 1267, 1274 (9th Cir. 1981), cert. denied, 455 U.S. 1021 (1982) (statistics rejected where the supposedly disadvantaged group had not prepared adequately for allegedly discriminatory test); Mozee v. American Commercial Marine Serv. Co., 940 F.2d 1036, 1047-48 (7th Cir. 1991) (claim of disparity rejected where statistics did not take into account prior disciplinary records of terminated individuals, which might have explained away disparity).

These pronouncements from the DOJ and federal courts around the country strongly support the premise of this memorandum: that bare statistical disparities in performance measures cannot constitute un rebuttable evidence of a BOC's alleged discrimination against competitors. Just as the defendant in an employment discrimination case must be given the opportunity to rebut the conclusions drawn by the plaintiff from the statistical disparities, so too must a BOC have the opportunity to explain apparent anomalies in reported performance measures. Due process requires no less.

#### IV. Constitutional Limits on the Use of Statistical Data

The Fifth Amendment's Due Process Clause provides that "[n]o person shall . . . be deprived of life, liberty or property, without due process of law." U.S. Const. amend. V. When a plaintiff has claimed that the procedures employed by the government fail to satisfy the Due Process Clause, a court engages in two inquiries. Cleveland Bd. of Educ. v. Loudermill, 470 U.S. 532, 541 (1985). First, the court asks whether the plaintiff has asserted a property interest that falls within the scope of the Clause. Second, the court determines whether the government accorded the plaintiff the process that was due. Id.

A protected property interest exists when there are "rules or understandings that secure certain benefits and that support claims of entitlement to those benefits." Board of Regents v. Roth, 408 U.S. 564, 577 (1972). This type of interest indisputably exists when a carrier asserts competitive rights that are established by the Telecommunications Act. Because the Act "secures certain benefits . . . [and] support[s] claims of entitlement to these benefits" for BOCs -- long distance relief -- a valid due process interest is at stake here.<sup>3</sup>

The FCC's use of un rebuttable statistical data to deprive a BOC of long distance relief or to impose monetary penalties violates the Fifth Amendment's guarantee of due process. This conclusion is supported by cases in which courts have recognized the due process implications that arise from the government's use of statistics to determine entitlements and obligations under government programs. For example, in Daytona Beach Gen. Hosp., Inc. v. Weinberger, 435 F. Supp. 891 (M.D. Fla. 1977), a federal agency used statistical sampling to determine the amount owed to the government by a Florida hospital under the Medicare program. The hospital challenged its repayment obligation on the ground that the use of sampling violated its due process rights. The district court agreed and entered summary judgment for the hospital.

Critical to the court's decision was the government's failure to give the hospital an adequate opportunity to challenge the statistics. The statistical sample the government selected for examination had been reviewed "without the participation of the plaintiff and without consultation with the attending physicians." Id. at 896. The use of the sampling procedure was discussed at only one of the six meetings held between personnel of the plaintiff and defendant. Id. at 896. Ultimately, the court

---

<sup>3</sup> And, of course, any potential monetary fine or penalty implicates a property interest when evaluating the constitutional limits on the use of statistical data in a self-executing penalty scheme.

agreed that when the government's use of statistics was coupled with "the failure to . . . provide[] a proper hearing . . . and the lack of adequate appeals procedure," a due process violation resulted. Id. See also Ratanasen v. California Dept. of Health Servs., 11 F.3d 1467, 1470-72 (9th Cir. 1993) (use of statistical sampling to determine a \$125,789 liability was lawful only because the doctor was given an opportunity to rebut the statistical evidence employed).

The same due process concerns arise when the FCC uses unrebuttable statistics to evaluate a BOC's section 271 application. The imposition of a penalty plan takes the procedural violation a step further because it gives the BOC no opportunity to respond.

## **V. Penalty Plans as Unrebuttable Presumptions**

The FCC has stated that "as a part of our public interest inquiry, we would want to inquire whether the BOC has agreed to private and self-executing enforcement mechanisms that are automatically triggered by noncompliance with the applicable performance standard without resort to lengthy regulatory or judicial intervention." Ameritech Michigan Order ¶ 394. According to the FCC, "[t]he absence of such enforcement mechanisms could significantly delay the development of local exchange competition by forcing new entrants to engage in protracted and contentious legal proceedings to enforce their contractual and statutory rights to obtain necessary inputs from the [BOC]." Id.

Similarly, in BellSouth Louisiana II, the FCC stated that "we would be particularly interested in whether such performance monitoring includes appropriate self-executing enforcement mechanisms that are sufficient to ensure compliance with the established performance standards." BellSouth Louisiana II Order ¶ 364.

Just as the FCC may not penalize U S WEST in section 271 proceedings by assuming that perceived disparities in statistical data are conclusive proof of discrimination, the FCC likewise may not insist on penalty plans which, after section 271 approval, create liquidated damages or penalties based upon the same types of statistical disparities. Even those who support penalty plans recognize that statistical disparities may not trigger automatic penalties. Tim Sloan, Creating Better Incentives Through Regulation: Section 271 of the Communications Act of 1934 and the Promotion of Local Exchange Competition, 50 Fed. Comm. L. J. 309, 381 (1998). Instead, if a disparity occurs, any issue regarding the BOC's compliance with the performance measure should be resolved expeditiously, through the use of binding arbitration or expedited agency procedures, but not without due process. Id.

Moreover, from the perspective of constitutional due process, it makes no difference whether the FCC seeks to impose a penalty plan as part of its "public interest" inquiry, or instead insists that the BOC's interconnection agreements contain liquidated damages provisions which have the same effect. By imposing "self-executing" penalties of any kind, the FCC violates due process just as much as does the failure to permit a defendant to rebut statistical evidence of employment discrimination. The FCC may not -- through a penalty plan or otherwise -- deprive the BOC of the opportunity to respond to statistical disparities.

## **VI. Conclusion Regarding the Use of Statistical Data**

The FCC should recognize the limits that the Constitution places on its proposed use of penalty plans. Statistical triggers for penalties are not only bad policy, they are bad law. Just as important, the BOCs' right to rebut statistical evidence of discrimination does not undermine the pro-competitive goals of the Act. Indeed, the Act's goal of promoting competition strongly supports giving BOCs the right to explain statistical disparities that are not caused by discrimination.

## Purpose of Presentation

- Discuss Usefulness of Statistical Testing
- Discuss Appropriateness of Tests Used by US WEST
- Discuss Appropriateness of Statistical Methods Used by US WEST

## Purpose of Statistical Tests

- Statistical Tests Are a First Step in the Analysis of Whether US WEST is Providing Non-Discriminatory Service
- Identification of Areas of Possible Service Difference Which Require Further Investigation
  - Products
  - CLECs
  - Time Periods

## Example of The Issue to be Resolved

- Observation:
  - The Reported Mean Time Required to Install a DS1 for CLEC A is Higher Than the Reported Mean Time Required to Install a US WEST DS1
- Question:
  - Is the Difference Due to Random Variation or a Real Difference in Service Quality?

## Possible Explanations for Difference in Reported Mean Time to Install

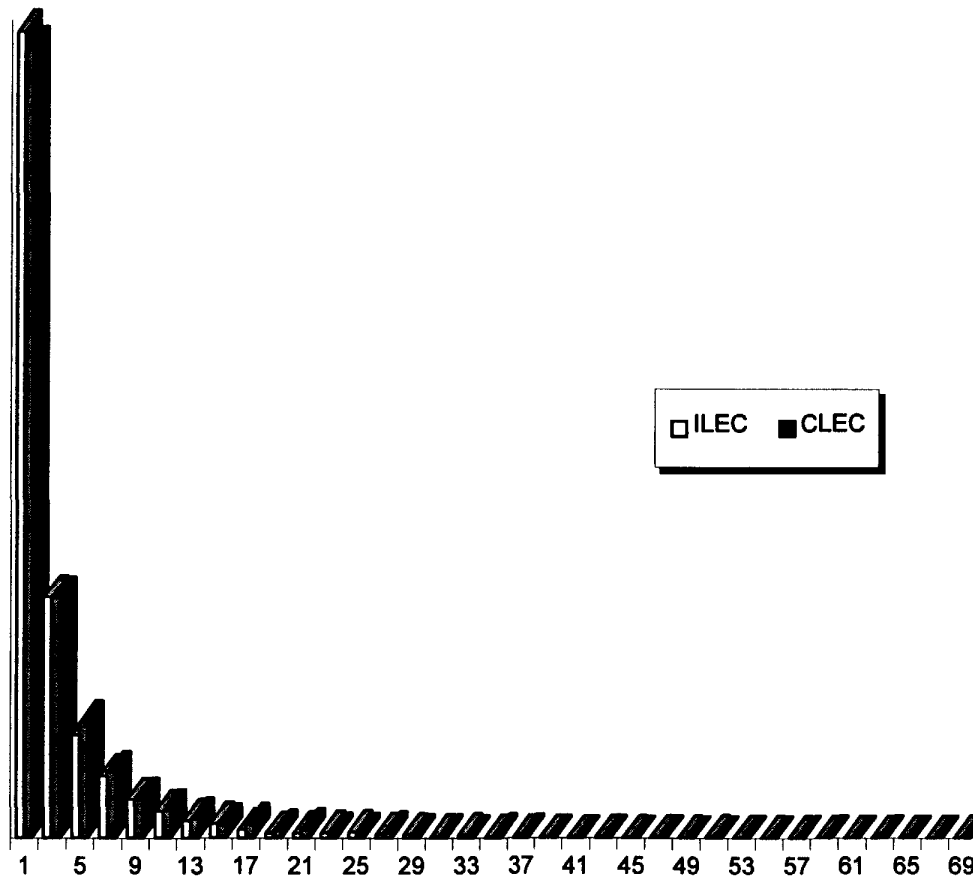
- CLEC A's DS1s are in Congested Locations
- CLEC A Orders Only Once Per Month
- All CLECs Order Only at End of Month
- CLEC A Orders in Large (Small) Batches
- CLEC A's Orders are Treated Differently
- All CLEC Orders are Treated Differently

## Other Explanations for Apparent Difference: Data Errors

- Technician Filled in Wrong Start Date
- Technician Forgot to Reset Work Code
- Technician Was Unaware of a Change to the Reporting Procedure
- Technicians Do Not Consistently Interpret the Applicable Reporting Procedure

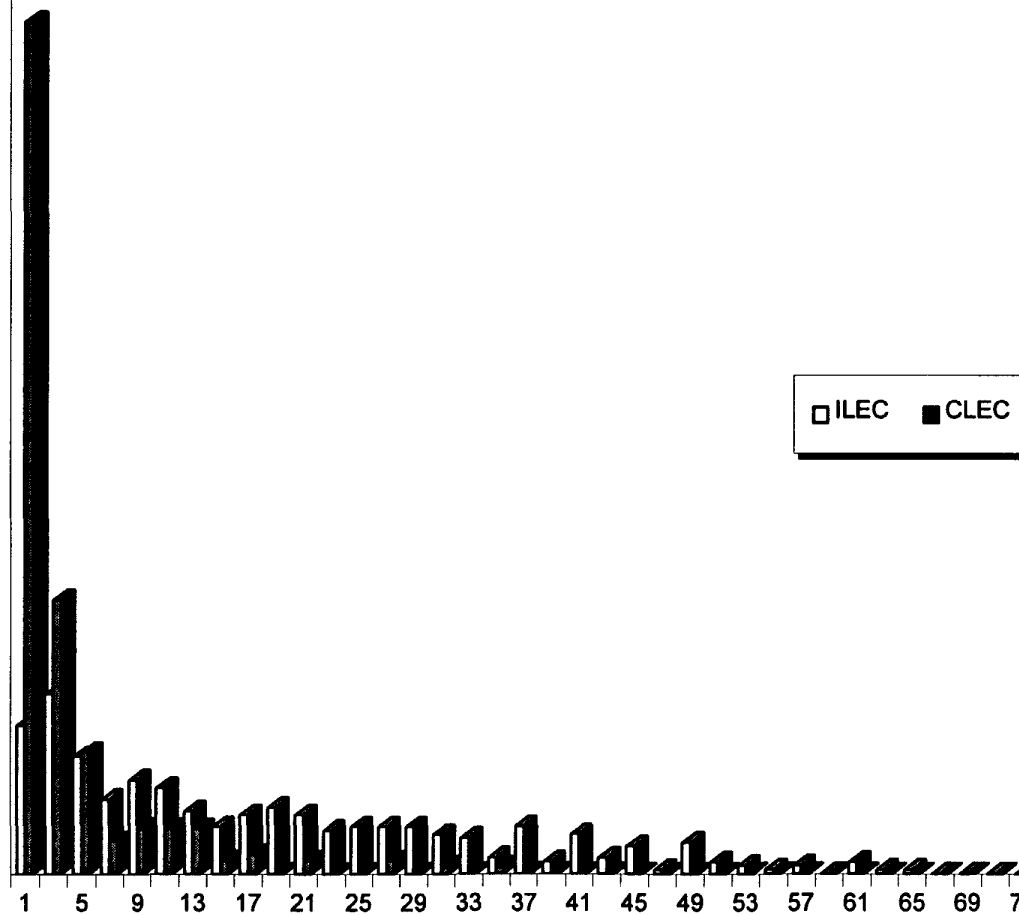
# Actual AZ DS1 Repair Times (CLEC & ILEC)

Seven Month Aggregate (Hours)



# Actual AZ DS1 Repair Times (CLEC & ILEC)

January 1999



# Overview of US WEST's Statistical Tests and Methods

- Aggregate CLECs vs. US WEST
- Individual CLEC vs. US WEST
- Individual CLEC vs. other CLECs
- Test for Adverse Difference in Mean/Proportion
- Permutation Test (CLEC Sample  $\leq 600$ )
- Z-Test (CLEC Sample  $> 600$ )
- 99 Percent Confidence (one tail)

# Why Is Permutation Test Required?

- Continuous Measures (means)
  - Non-Normal, Population Skewed Distributions
    - Must Rely on Large Samples For Normality
  - Unequal Sample Sizes
  - Small Sample Sizes
  - Monte Carlo Simulations Indicate That Using a Normal Assumptions Will Result in Lower Than Specified Confidence Levels (Probability of a False Positive is Greater Than One Minus the Confidence Level)

# Why Is Permutation Test Required?

- Proportion/Rate Measures
  - Proportions can be close to zero or one
  - Unequal sample sizes
  - Small sample sizes
  - Monte Carlo simulations indicate that the Normal approximation to the binomial distribution can result in higher than specified confidence level

## Hypotheses of a Permutation Test

- If both samples are drawn from the same population (null hypothesis) the observed value of the test statistic will fall within the “center” of the distribution of all possible values that could be generated with the same pool of data.
- If the observed or even less favorable values result from only very few of the possible rearrangements of the pooled data, it is improbable that the samples are drawn from the same population.
- Therefore they must originate from different populations (alternate hypothesis)

## Advantages of a Permutation Test

- Can be Used With Smaller Samples Because It Does Not Rely On the Central Limit Theorem to Assure Normality of Distributions
- Allows Use of a “Modified” Test Statistic Which is More Sensitive to Differences Which May More Adversely Affect CLECs Does Not Alter Confidence Level
- Constructs empirical distribution of statistic from permutations of the actual observations

## Limitations of Permutation Test

- Does Not Address All Inherent Limitations of Statistical Tests
  - Systematic variation, independence of samples
  - Data “Cleanliness” (A minimum sample size of 20 reduces the possibility that a single erroneous data point will cause a test failure)
- Computationally Intensive
  - Opportunities for mitigation
- Individual Results Are Not Strictly Reproducible ( A random sample of permutations is used for continuous measures)

## Z-Test is Adequate for Very Large Sample Sizes

- For CLEC Samples Larger than 600 the Standard Z-Test Provides Results Consistent with the Permutation Test
- For Continuous Measures:

$$Z = \frac{(\bar{X}_C - \bar{X}_I)}{\sqrt{\frac{\sigma_I^2}{N_I} + \frac{\sigma_C^2}{N_C}}}$$

## Z-Test is Adequate for Very Large Sample Sizes

- For Proportional Measures The Binomial Distribution is Closely Approximated by a Normal Distribution:

$$Z = \frac{(p_C - p_I)}{\sigma_D}$$

$$\sigma_D = \sqrt{\frac{p(1-p)}{n_I + n_C}}$$

$$\text{and } p = \frac{x_I + x_C}{n_I + n_C}$$

## Importance of Selecting the Appropriate Confidence Level

- Probability of False Positive is Solely Determined by Selection of Confidence Level
- Where Consequences of Test “Failure” Are Serious, the Probability of False Positive Must be Low
- Monetary Penalties Require Low Probability of False Positives
  - If not, even absolutely equal service will simply transfer substantial funds to CLECs
- Integrity of the Test Dictates a High Confidence Level

## US WEST's Confidence Level

- USWEST Uses 99 Percent Confidence Level to Minimize False Positives
  - If Sample Sizes are Large, Low Confidence Level Increases Failures, but the Additional Failures will be Primarily False Positives
  - Single Tests Are Not Provided in Isolation, Both Aggregate and Individual Comparisons Are Provided as Well as Comparisons to All Other CLECs
  - Prudence and avoidance of the severe consequences of a failure at high confidence should trigger an investigation upon failure at lower confidence

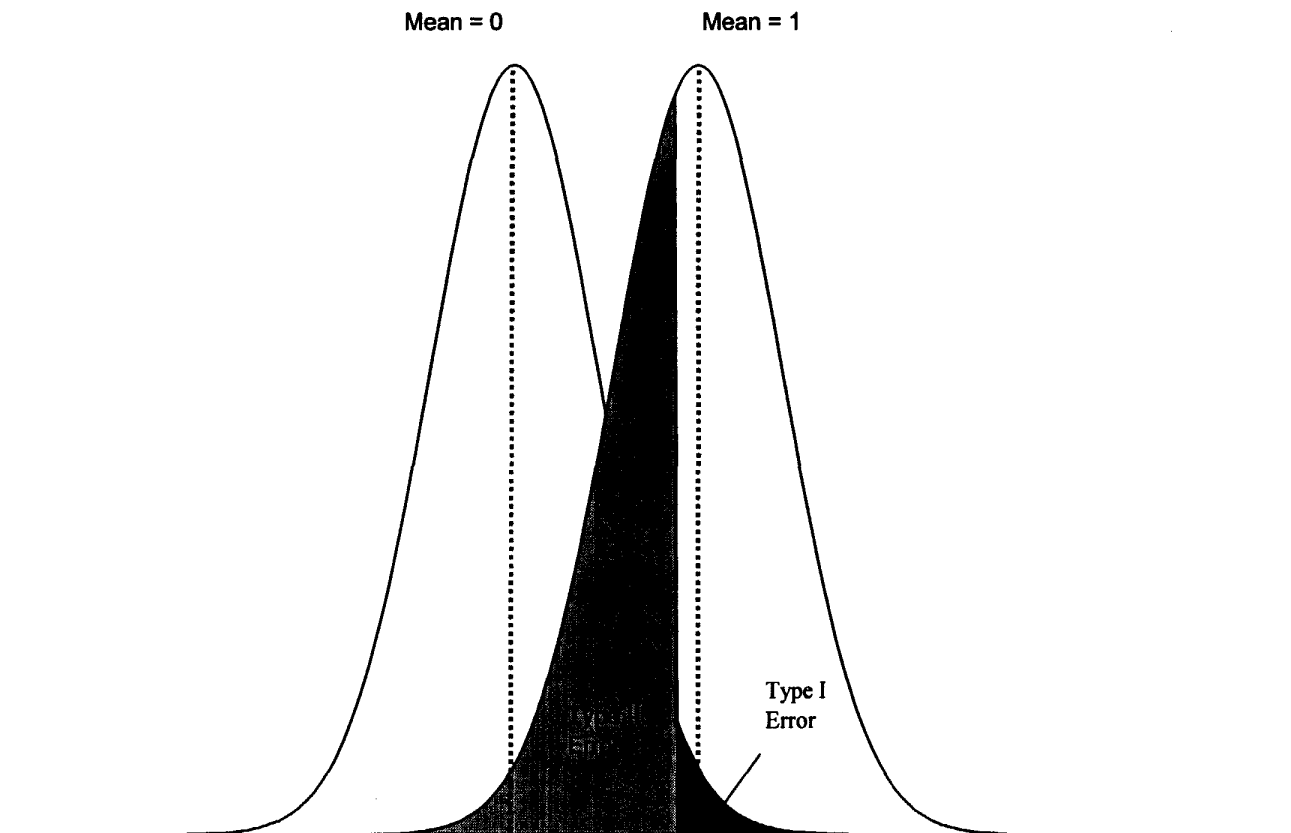
## Confidence Level: Error Types and Their Probabilities

- Type I: Declaring a Difference When There is Only Random Variation (False Positive)
- Type II: Failure to Detect a Difference in Service Quality When It Exists
- Most Desirable Test Has a Low Probability of Both Type I and Type II Error

## Confidence Level: Factors Which Determine Error Probabilities

- Type I: Determined Solely by Confidence Level
- Type II: Depends on Several Factors
  - Confidence level
  - Variance of population (samples)
  - Size of samples
  - Magnitude of difference to be detected

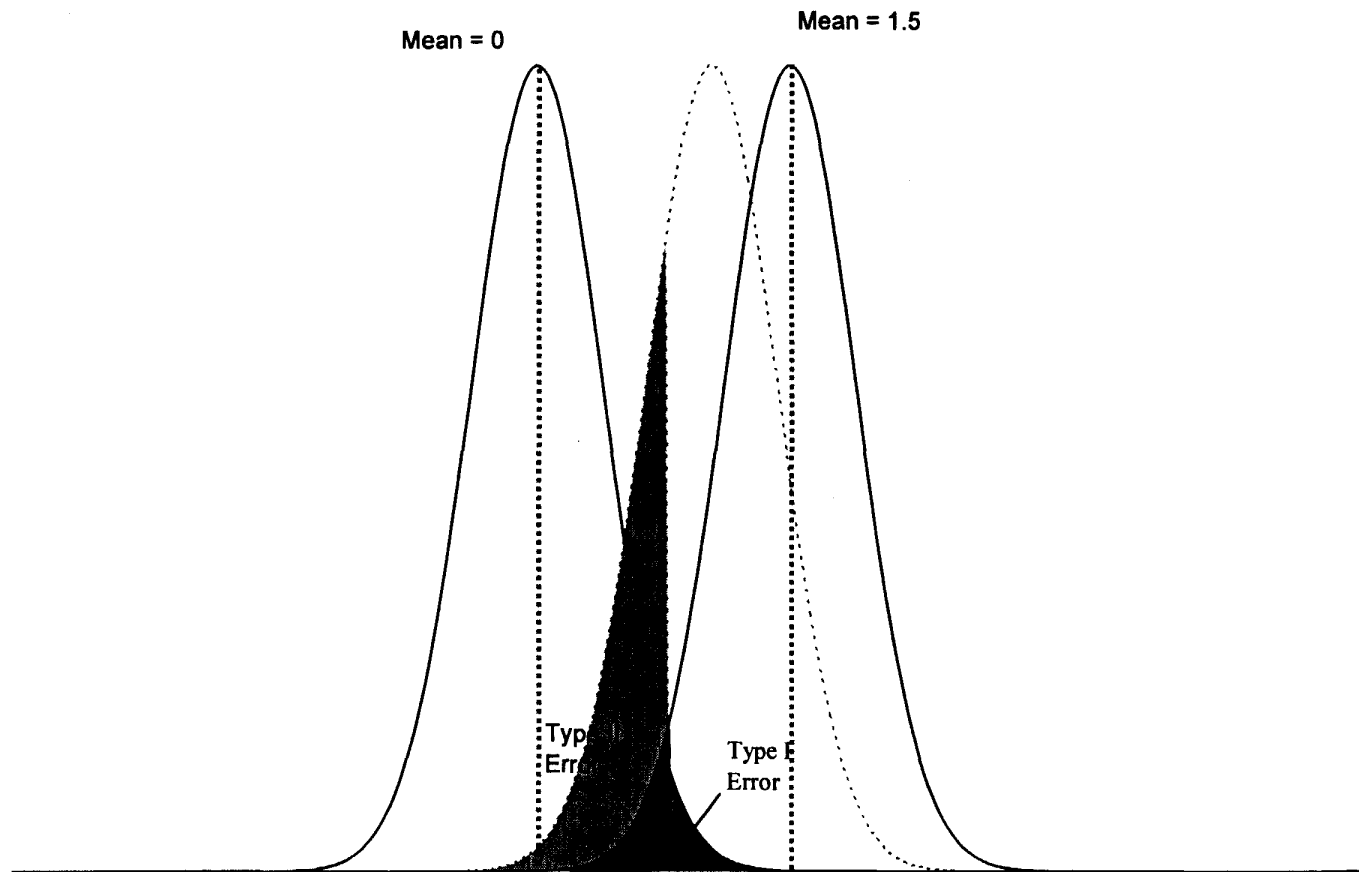
# Error Types and Their Probabilities



## Confidence Level: Reduction of Type II Error

- Type II Error Probabilities Can be Reduced Without Recourse to Low Confidence Levels
  - Aggregate CLEC results will show systematic differences with lower Type II Error
  - Aggregation over time can be used to decrease Type II Error
  - If there is no systematic difference between CLEC & ILEC operations, appropriately aggregated data provides a valid test of difference in service

# Realistic Difference in Means



## Appropriate Confidence Level

- The Need for a Low Probability of False Positives Dictates the Use of a High (99%) Confidence Level
- Type II Error Can be Controlled Through Use of Aggregate Monthly Measures
- Aggregation Over Time Can be Used to Further Reduce Type II Error

## Conclusion

- Statistical Tests Provide Substantial Information to Help Answer the Question of Whether There is a Real Difference in Service
- Characteristics of the Data Require a Permutation Test
- Serious Consequences of Failure Require High Confidence Level
- Nature of the Data Indicates the Specification of a Minimum Sample Size

End of Presentation

## Use of Absolute Rather Than Relative Standard

- Absolute Standards Will not Reflect:
  - Changes in Conditions
    - Seasonal weather or demand
    - Unusual Peak Demands (Ames effect)
  - Overall Increase in “Quality”
    - Technology changes could leave CLECs behind
  - Can Lead to Deterioration of Service
    - ILEC may be required to favor CLEC customers
  - Customer Demand for Lower Cost, Lower Quality Products
    - Restricted but Inexpensive Airline Tickets

# Appendix

## Implementation of Permutation Test

- Continuous Measures:
  - Construct 1,000 random permutations
  - Samples of size  $N_C$  and  $N_I$
- “Modified Z” Test Statistic:

$$Z = \frac{(\bar{X}_C - \bar{X}_I)}{\sqrt{\frac{\sigma_I^2}{N_I} + \frac{\sigma_I^2}{N_C}}}$$

## Implementation of Permutation Test

- If the Observed Value of the Test Statistic is Greater than the 11th Highest of the 1,000 Values, the Test Does Not Support the Hypothesis That The Samples are Drawn From the Same Population

000000 000000 000000  
000000 000000 000000

## Implementation of Permutation Test

- Proportional Measures:
  - Calculate the probability of finding a difference less favorable than that actually observed

$$P_D = \sum_{i \in D} \sum_{j \in D} P_{BIN}(i, p, N_C) P_{BIN}(j, p, N_I)$$

$$\text{and } p = \frac{x_I + x_C}{n_I + n_C}$$

## Values Reported

- Numerators
- Denominators
- Results (mean or proportion)
- Standard Deviations
- Comparison Standard Deviation
- Statistical Score

## Comparison Standard Deviation

- Continuous Measures

$$\sigma_C = \sqrt{\frac{\sigma_I^2}{N_I} + \frac{\sigma_I^2}{N_C}}$$

## Comparison Standard Deviation

- Proportional Measures

$$\sigma_D = \sqrt{\frac{p(1-p)}{n_I + n_C}}$$

$$\text{and } p = \frac{x_I + x_C}{n_I + n_C}$$

## Statistical Score

- Indicates Relative Distance Above or Below the Critical Value
- Positive Values Indicate Difference is Significant at the 99% Level

$$\text{Score} = \frac{\text{Actual Value} - \text{Critical Value}}{\text{Critical Value}}$$